

EFFECTOS DE UN INSTRUMENTO PARA EVALUAR PROYECTOS
DE INVESTIGACIÓN EXPERIMENTAL

por

ENRIQUE MARTÍNEZ OROZCO

TRABAJO DE INVESTIGACIÓN

Presentado a la Facultad de Educación de la
Universidad del Valle de Guatemala
como requisito previo a
optar el Grado de

MAESTRÍA

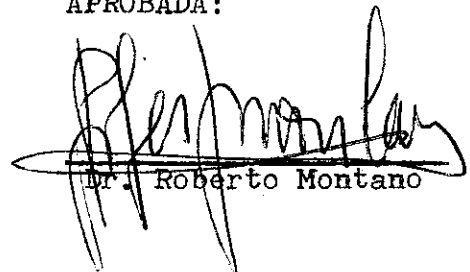
en

MEDICIÓN, EVALUACIÓN E INVESTIGACIÓN EDUCATIVA

Guatemala, noviembre, 1974

EFFECTOS DE UN INSTRUMENTO PARA EVALUAR PROYECTOS
DE INVESTIGACIÓN EXPERIMENTAL

APROBADA:



Mr. Roberto Montano

ÍNDICE

Capítulo		Página
I	INTRODUCCIÓN	1
II	REVISIÓN DE LITERATURA	4
III	HIPÓTESIS Y DISEÑO EXPERIMENTAL	15
IV	METODOLOGÍA	18
V	DISCUSIÓN, CONCLUSIONES, LIMITACIONES Y RECOMENDACIONES	44
	ANEXOS	51
	REFERENCIAS BIBLIOGRÁFICAS	83

CAPÍTULO I

INTRODUCCIÓN

En el curso de Maestría en Medición, Evaluación e Investigación Educativa de la Universidad del Valle de Guatemala, se incluye una materia llamada Diseño experimental cuyos objetivos principales son: primero, que el estudiante obtenga los recursos necesarios para iniciarse en la investigación y en particular para presentar una propuesta formal de un trabajo de investigación; segundo, que el estudiante adquiera con respecto a los informes de investigación educativa no sólo capacidad crítica para su lectura, sino también habilidad para redactarlos y organizarlos.

Como prueba final de la obtención de algunos de los objetivos del curso por parte de los alumnos, es necesario presentar un trabajo de investigación, que es requisito para obtener el grado de Maestría.

El trabajo de investigación incluye la elaboración de un proyecto del mismo, como el estudiante es novicio en investigación, toda ayuda con respecto a una estructuración y sistematización de la elaboración del proyecto redundará en un ahorro de tiempo por un lado y en una mayor calidad por el otro. Una forma donde el estudiante puede lograr esta estructuración y sistematización es a través de un instrumento que le permita seguir un procedimiento adecuado y además autoevaluar cada uno de

los pasos que constituyen el proyecto de investigación. Además con este instrumento el estudiante puede lograr una discriminación adecuada de aquella literatura científica que tiene cierta calidad para fundamentar su trabajo.

En la evaluación posterior del proyecto de investigación por parte del profesor del curso, se hace necesario la sistematización de la experiencia del evaluador a través de un instrumento, con el objeto de que éste evalúe con cierta homogeneidad de criterio los diferentes proyectos de investigación presentados por los alumnos.

Si bien existen estudios sobre el uso de instrumentos para la evaluación de informes de investigación experimental, estos han sido en su mayoría realizados por expertos estadounidenses. Como la ciencia y la tecnología está invadiendo todos los países, en particular los de latinoamérica, se considera necesario la creación y/o adaptación de instrumentos de evaluación de informes de investigación a este medio.

El propósito de este trabajo fue estudiar el efecto de un instrumento de evaluación de proyectos de investigación experimental utilizado por alumnos inscritos en un curso de posgrado en Medición, Evaluación e Investigación Educativa de la Universidad del Valle de Guatemala.

Definiciones

Proyecto--propuesta formal de una investigación experimental.

Instrumento--significa la adaptación del instrumento de Suydam (1968) la cual consiste de siete pautas de evaluación de proyectos de investigación experimental educativa y una guía del instrumento.

Evaluador--cada uno de los estudiantes de la Maestría en Medición, Evaluación e Investigación Educativa perteneciente a la muestra seleccionada en el presente estudio.

CAPÍTULO II

REVISIÓN DE LITERATURA

Este capítulo tiene por objeto incluir inicialmente, aquella literatura científica pertinente al presente trabajo, la cual se ha distribuido para su presentación en la forma siguiente; primero, juicios sobre proyectos de investigación; segundo, juicios sobre informes de investigación; tercero, estudio relacionado con la aplicación de un instrumento de evaluación de informes de investigación educativa experimental y cuarto, estudio sobre una guía para la evaluación de informes de investigación. Finalmente, se dan las razones para configurar el instrumento de evaluación de proyectos de investigación experimental educativa, propuesto en el presente estudio, en la forma en que se hizo.

Juicios sobre proyectos

Ary y otros (1972) enfatizan que escribir el proyecto de investigación es la etapa más crucial y excitante en el proceso de investigación; en esta etapa el proceso cristaliza en forma concreta. En el proyecto el investigador demuestra que sabe lo que está buscando, como lo verá si está allí, y porqué la investigación vale la pena. Un proyecto bien definido por parte del investigador indica que él está realmente listo para poner su estudio en movimiento y muestra que sabe lo que hace, por qué lo hace, y como hacerlo. Si el investigador no puede

producir un proyecto completo y bien definido, no estará aún en condiciones de proceder a la etapa de recolección de datos de su investigación. Ary y otros, mencionaron también que "recolectar los datos y escribir el proyecto simultáneamente puede parecer un procedimiento sano de ahorrar tiempo pero tal caso es raro; innumerables horas-hombre y miles de dólares han sido desperdiciados justamente de esta manera (pág. 233)". Hasta que el proyecto está formulado, no se puede asegurar qué datos serán necesarios, ni cuál será la mejor forma de manejar esta información a la luz del propósito del estudio. Por otra parte, cuando el investigador trabaja en base a fechas fijadas de antemano para cumplir ciertas etapas de la investigación, debe establecer una fecha para completar el proyecto antes del día fijado para terminar la investigación total.

En general, dice Ary y otros, el investigador necesita presentar proyectos en forma escrita y organizada. Esta presentación consta de dos etapas: el proyecto de investigación y el reporte de los resultados de la investigación. En la primera etapa deben aparecer los siguientes rubros: establecimiento del problema, revisión de la literatura, pregunta y/o hipótesis, instrumentación, muestreo, diseño, recolección de datos y procedimientos estadísticos, nivel de significación, presupuesto y cédula de tiempo. Esta lista proporciona una útil guía para la redacción de cualquier proyecto.

Best (1969) señala que la preparación de un proyecto es un paso importante en el proceso de la investigación. Muchas

instituciones exigen que un proyecto sea presentado a su consideración antes de aprobarlo. Este proyecto proporciona un fundamento para su evaluación y da una base para aconsejar durante el período que dure la investigación, también proporciona un plan sistemático que ha de seguir el investigador. Best propone asimismo, como las partes esenciales de un proyecto de investigación: el enunciado del problema, la importancia del problema, definiciones, supuestos y limitaciones, un resumen de la bibliografía acerca del tema, un cuidadoso y detallado análisis de los métodos de investigación propuestos, y un esquema de tiempo. El proyecto puede ser muy modesto y necesariamente limitado por el tiempo, como puede ocurrir en el caso de la necesidad de un estudiante de presentar un proyecto de investigación para la opción al grado universitario.

Cook (1966) dice que es difícil enfatizar la importancia del proyecto de investigación como base para realizar una investigación. En el caso de un estudiante de un curso para principiantes en investigación educativa, la importancia radica en la presentación de un cuidadoso proyecto ya que, debido a su poca experiencia, el estudiante puede olvidar la inclusión de pautas importantes en la investigación que hagan ésta inválida. Entre mayor esfuerzo y cuidado pongan los estudiantes en la elaboración del proyecto, más probabilidad de éxito tendrá la ejecución de la investigación. Con respecto al contenido de un buen proyecto, Cook dice que no existe una respuesta directa a esta pregunta, y cualquier sugerencia dependerá, en alguna

extensión, sobre las condiciones en las cuales la investigación se hace. La clase de proyectos que un estudiante principiante somete antes de intentar un proyecto de investigación para un curso de métodos de investigación, debe tener probablemente seis partes principales: título tentativo, establecimiento del problema, método de investigación y procedimiento, datos a ser recolectados, de dónde serán obtenidos los datos e investigación relevante y teoría.

Según Hayman (1969), antes de comenzar a trabajar activamente en una investigación se debe redactar un plan en detalle. A menudo esta redacción tomará la forma de un proyecto pero, cualquiera que sea la forma que se le dé, es muy importante formular un plan completo de la investigación antes de comenzar el trabajo efectivo. Un error que cometen comúnmente los investigadores, en especial los inexpertos, consiste en comenzar un estudio sin un adecuado planteamiento previo, y por lo tanto, el investigador comienza a reunir una serie de evidencias dispersas que no tienen ninguna relación con su estudio y que además le hacen perder tiempo valioso para la realización de la investigación.

Hayman dice que los pasos que debe incluir un plan o proyecto son: selección de un problema, formulación por escrito del problema, desarrollo de una teoría apropiada, formulación de las hipótesis y mediante ellas, formulación de los objetivos específicos del estudio, planeamiento de los procedimientos a

seguir en la realización del estudio y en la selección de los métodos de investigación que se utilizarán, los datos necesarios y los medios destinados a obtenerlos, las técnicas a utilizar en el análisis de los datos obtenidos y la muestra de estudio, así como el método para seleccionarla.

Juicios sobre informes

Después de aceptado el proyecto, la siguiente etapa en la investigación es la que constituye propiamente la parte dinámica de la investigación que comienza con la recolección de datos y termina con el informe de la investigación.

Best (1969) propone un esquema de evaluación de un informe de investigación constituido por los siguientes grandes rubros y sus correspondientes subrubros: problema, significación, planteamiento, delimitación, hipótesis específicas, supuestos establecidos, limitaciones reconocidas; desarrollo del problema, apropiado, descripción detallada, planeamiento cuidadoso, recolección sistemática de datos; análisis y conclusiones, lógica, basados sobre los datos presentados con objetividad, resumidos con efectividad; bibliografía, revisión adecuada, bien organizada; forma y estilo, márgenes, colocación, espaciado, epígrafes, notas, citas, paginación, mecanografía, tablas, figuras, bibliografía, ortografía, redacción, estilo, otras características.

Zuñiga (1967) propone las siguientes pautas de evaluación de informes de investigación, con sus correspondientes

subrubros que las integran: formulación del problema, antecedentes, justificación, definición, alcance y límites; marco conceptual, premisas, hipótesis, definición de términos; fuentes y métodos, procedimiento general; muestreo; pruebas, contenido de las pruebas, prueba preliminar; cuestionarios, forma de las preguntas, orden de las preguntas, respuestas; clasificación de los datos; recopilación de los datos; análisis de los datos; conclusiones; algunas anotaciones sobre estudios de tipo histórico; consideraciones sobre estudios de tipo descriptivo; consideraciones sobre estudios de tipo experimental; consideraciones sobre organización y formato.

Fox (1969) propone la división del proceso de investigación en tres partes: primero, diseño del plan de investigación; segundo, implementación del plan de investigación y tercero, implementación de los resultados. La primera parte incluye: idea inicial y área del problema, revisión inicial de la literatura, definición específica del problema a investigar, estimación de los sucesos potenciales de la investigación contemplada, segunda revisión de la literatura, selección del abordaje de la investigación, establecimiento de las hipótesis de la investigación, selección de métodos y técnicas de recolección de datos, selección y desarrollo de los instrumentos de recolección de datos, diseño del plan de recolección de datos, identificación de la población y muestra considerada, estudios para abordar la recolección de datos. La segunda parte incluye: implementación del plan de recolección de datos, implementación del plan de

análisis de datos, preparación del reporte de investigación. La tercera parte está formada por la difusión de los resultados y agitación para la acción.

Estudio de un instrumento de evaluación de informes

Suydam (1968) diseñó un instrumento para evaluar informes de investigación experimental educativa a partir de una síntesis de las sugerencias hechas por algunos expertos. En una recopilación hecha por él a partir de las listas de las sugerencias hechas por los expertos en el campo de la evaluación de investigación educativa, encontró que nueve pautas eran repetidas consistentemente y éstas formaron la base de su instrumento para evaluar informes de investigación experimental educativa. Estas nueve pautas son: la significación teórica y práctica del problema, la definición del problema, la adecuación del diseño y las preguntas de la investigación, el control de las variables por el diseño, la adecuación de la muestra seleccionada para el diseño y el propósito de la investigación, la validez y confiabilidad de los instrumentos o técnicas empleadas, la validez de las técnicas de análisis de datos, la interpretación y generalización de los datos y la adecuación del reporte de investigación.

Suydam hizo dos estudios con el objeto de determinar la confiabilidad o grado de acuerdo entre evaluadores que emplearon el instrumento. El primer trabajo en particular tuvo por objeto estudiar la confiabilidad del uso del instrumento,

sin que dicha confiabilidad se viera aumentada por el entrenamiento previo de los evaluadores y para ello el período de entrenamiento fue limitado. Entre los artículos de investigación experimental seleccionados de los publicados en The American Teacher en el período de 1954 a 1965, escogió una muestra al azar de diez de dichos artículos para su reproducción y evaluación. Dentro de esta muestra seleccionada al azar, a su vez escogió al azar cinco artículos a los cuales les suprimió el nombre del autor y el año de la publicación, de tal forma de hacer posible el testar el efecto de la condición del autor y del año de publicación. Los evaluadores, en número de tres, eran candidatos al doctorado en educación elemental. Ellos efectuaron las evaluaciones de los diez artículos independientemente, usando el instrumento de evaluación de informes de investigación propuesto por Suydam, encontrando éste, un coeficiente de correlación de .91 utilizando la fórmula de análisis de varianza para obtener el coeficiente de correlación entre evaluadores.

Para medir la probable consistencia del uso del instrumento por parte de un mismo evaluador se utilizó la fórmula de Snedecor para la confiabilidad de evaluaciones individuales. El coeficiente de confiabilidad encontrado con el uso de esta fórmula fue de .77.

En el segundo trabajo, Suydam, tuvo por objeto medir la utilidad del instrumento para diversos tipos de profesionistas ligados con la evaluación de investigación educativa, y a quienes

supuestamente sólo es posible darles un entrenamiento limitado o nulo sobre dicho uso del instrumento, para semejar esta condición en el estudio Suydam sólo consideró como orientación para los evaluadores las instrucciones dadas en el instrumento. Entre artículos de investigación experimental, tomados de los publicados en revistas de los Estados Unidos en el período de 1930 a 1965, Suydam seleccionó una muestra estratificada de diez de esos artículos, basando su estratificación en la fuente de la revista, la condición del autor y el año de la publicación. Con el objeto de establecer el efecto de la condición del autor, la fuente de la revista y el año de la publicación, estos datos fueron suprimidos en cinco de los diez artículos que forman la muestra aleatoria de artículos a evaluar, los cinco artículos fueron seleccionados al azar. Los evaluadores considerados en el estudio fueron doce distribuidos en la forma siguiente: tres candidatos al doctorado en educación elemental, tres candidatos al doctorado en psicología educativa y tres profesores en psicología educativa, teniendo como común denominador, estos grupos de evaluadores, el ser representativos de sectores interesados en evaluación de investigación educativa. La evaluación de los diez artículos fue hecha independientemente por los evaluadores aplicando el instrumento de evaluación de informes de investigación experimental propuesto por Suydam, encontrando éste, un coeficiente de correlación entre evaluadores de .94 utilizando la fórmula de análisis de varianza para obtener el coeficiente de correlación entre evaluadores. Con la

fórmula de Snedecor para la confiabilidad de evaluaciones individuales el coeficiente de confiabilidad encontrado fue de .57.

Estudio sobre una guía de evaluación de informes de investigación

Strauss (1969) propuso un conjunto de lineamientos para analizar informes de investigación, principalmente de investigación experimental. Estos lineamientos consisten en veinte criterios que fueron probados durante varios años en estudiantes que no habían tenido entrenamiento formal en métodos de investigación por un lado y en algunas otras personas que habían tenido entrenamiento en diferentes disciplinas de la educación. Los criterios probados son relativamente obvios y simples de aplicar y relacionan algunos hechos del informe de investigación que son descuidados algunas veces. Los criterios propuestos por Strauss son: establecimiento del problema, trabajos previos citados, objetivos establecidos, hipótesis formuladas, presuposiciones hechas, población estudiada, extracción de la muestra, instrumentos usados, examen del diseño, procedimiento seguido, seguridades tomadas, observaciones registradas, resultados registrados, interpretación estadística, discusión de interpretaciones, conclusiones alcanzadas, limitaciones reconocidas, proyección de otros trabajos, mejoramientos sugeridos y claridad del informe.

Sumario

Los estudios anteriores permiten fundamentar la elección de un mínimo de siete pautas de evaluación para un

instrumento de evaluación de proyectos de investigación experimental. También permiten considerar que un informe de investigación completo se puede dividir en dos partes, una es el proyecto de investigación y la otra es el informe de investigación. Las siete pautas que han sido seleccionadas, tomando en cuenta fundamentalmente el trabajo de Suydam, coinciden con las pautas seleccionadas también por otros investigadores así como con las pautas mencionadas en la revisión de literatura.

CAPÍTULO III

HIPÓTESIS Y DISEÑO EXPERIMENTAL

En el capítulo anterior se discutió el marco teórico del presente trabajo, en el presente capítulo se da introductoriamente las hipótesis convencionales a ser testadas en el diseño experimental, a continuación se presentan las variables relevantes de dicho diseño y finalmente se informa al lector de los modelos estadísticos utilizados para testar las hipótesis bajo estudio.

Hipótesis

Para poder analizar los efectos de adaptación por el autor del instrumento diseñado por Suydam se formularon las siguientes hipótesis:

Hipótesis uno: la media de los puntajes obtenidos por los evaluadores que aplicaron el instrumento de evaluación de proyectos de investigación experimental es igual a la media de los puntajes obtenidos por los evaluadores que no aplicaron el instrumento.

El valor crítico para la confirmación de la hipótesis fue establecido a un nivel de significación de 0.05.

Hipótesis dos: la desviación estándar de los puntajes obtenidos por los evaluadores que aplicaron el instrumento de

evaluación de proyectos de investigación experimental es igual a la desviación estándar de los puntajes obtenidos por los evaluadores que no aplicaron el instrumento.

El valor crítico para la confirmación de la hipótesis fue establecido a un nivel de significación de 0.05.

Diseño experimental

Este subtítulo comprende la descripción de las variables relevantes del estudio en un principio para posteriormente también describir como fueron testadas las hipótesis convencionales sujetas a estudio.

La variable independiente--es la evaluación del trabajo y constará de dos niveles: el primer nivel será la evaluación con instrumento y el segundo la evaluación sin instrumento.

La variable dependiente--los puntajes dados por los evaluadores.

Las constantes--el nivel del curso, la escuela, el profesor del curso.

Grupo experimental--tomando como base la premisa de que el conocimiento del vocabulario ha sido frecuentemente usado como un índice del conocimiento de la persona de un campo especializado Ebel (1960), se aplicaron a los integrantes de la población bajo estudio dos inventarios: uno de vocabulario

básico en medición y evaluación educativa Ebel (1960) traducido por el investigador y agregada una lista de los términos del vocabulario en español y otro de vocabulario básico en investigación educativa elaborado por el investigador, completado con una lista de los términos del vocabulario en inglés. Para el estudio, los sujetos fueron asignados al azar en dos grupos cada uno de los cuales contribuyó al establecimiento de los niveles de la variable independiente como se describirá en la selección de la muestra.

El modelo estadístico--para el testado de la hipótesis uno se usó el modelo "t" de Student, el cual fue seleccionado debido a que la muestra empleada fue menor que 30 y por lo tanto se hizo necesario utilizar una distribución muestral de estadísticos para pequeñas muestras como lo señala la teoría estadística. El test hecho fue de dos colas.

Para testar la hipótesis dos, la probable diferencia en variabilidad, medida en desviación estándar, entre los puntajes obtenidos por los evaluadores que utilizaron el instrumento y los que no lo emplearon, se hizo una prueba F de significación de la diferencia entre desviaciones estándar en muestras pequeñas; esta prueba fue seleccionada debido a que la muestra considerada es menor de 30, muestra pequeña, y no correlacionadas.

El test aplicado fue bilateral y a un nivel de significación de 0.05. Los grados de libertad de este test fue de $N-1$ para ambos puntajes debido a que el número de observaciones es igual para ambos grupos de evaluadores.

CAPÍTULO IV

METODOLOGÍA

Se presentaron en el capítulo anterior las hipótesis convencionales y el diseño estadístico. En este capítulo con el objeto de describir el análisis de datos como punto final de una metodología empleada, se hace necesario proceder a definir primeramente la población bajo estudio, para inmediatamente realizar la selección de la muestra. A continuación se describe la selección de materiales usados por todos los evaluadores en el experimento así como el instrumento de evaluación empleado por los evaluadores del grupo experimental y los instrumentos de medición del conocimiento de vocabulario básico de medición y evaluación educativa y de medición del conocimiento de vocabulario básico de investigación educativa. Se termina este capítulo con el análisis de datos precedido de la descripción de la recolección de datos pertinentes al estudio.

Población

La población sobre la cual se hizo este estudio, estuvo constituida por un total de 16 alumnos de la Maestría en Medición, Evaluación e Investigación Educativa de la Universidad del Valle de Guatemala que cursaron sus estudios durante el año académico de 1974. Esta población se encontró constituida por profesionales en las siguientes ramas: nueve pedagogos, dos

psicólogos, dos licenciados en filosofía y letras, una economista, un odontólogo y un matemático; todos estos profesionales tienen experiencia en la docencia que varía desde un mínimo de dos años hasta un máximo de 15. La edad promedio de la población fue de 29 años y estuvo constituida por cuatro mujeres y doce hombres.

Selección de la muestra

La muestra para este estudio se obtuvo haciendo una lista de los integrantes de la población descrita anteriormente asignándoles a cada uno un número y utilizando una tabla de números aleatorios se seleccionó seis integrantes de la población. De estos seis se decidió, por medio de una moneda, la mitad que utilizaría el instrumento y la mitad que no. Uno de los sujetos seleccionados tuvo que ser sustituido arbitrariamente por otro por causas ajenas a este investigador.

Selección de materiales

El primer problema con que se encontró el autor fue el escoger la literatura apropiada para ser evaluada en el presente estudio. La selección de los artículos a ser considerados en la investigación de que se informa, se hizo después de una cuidadosa revisión de los números existentes de la Revista Interamericana de Psicología a partir de 1969 a 1974 y de los números existentes de la Revista Mexicana de Psicología de 1963 a 1971, en la biblioteca de la Universidad del Valle de Guatemala.

De la revisión, del contenido de las revistas anteriormente mencionadas, se obtuvieron 15 artículos que correspondían a otros tantos informes de investigación experimental. Estos artículos seleccionados fueron numerados en una lista hecha a propósito y por medio de una tabla de números aleatorios se escogieron al azar tres artículos para ser reproducidos y evaluados posteriormente por la muestra de evaluadores seleccionados dentro de la población considerada para el estudio; estos artículos fueron: El "dibujo de la persona" en esquizofrénicos y normales y el MMPI (español) aplicado a jóvenes mexicanos: influencia de sexo, edad y nivel de inteligencia cuya fuente fue la Revista Interamericana de Psicología y de la Revista Mexicana de Psicología el artículo Influencia del prejuicio del experimentador sobre la observación de un evento físico.

Descripción del instrumento

En este subtítulo se describirán los tres instrumentos empleados en este estudio; el primero que se presentará será el instrumento propuesto para evaluar proyectos de investigación experimental educativa y finalmente dos instrumentos más que corresponden a dos inventarios, uno de conocimiento de vocabulario básico en medición y evaluación educativa y otro de vocabulario básico en investigación educativa.

Instrumento propuesto para evaluar proyectos--el instrumento propuesto para evaluar proyectos de investigación experimental educativa (ver Anexo A) es, en parte, una adaptación

del instrumento de evaluación de informes de investigación experimental propuesto por Suydam (1968), al cual se le anexó una guía para el instrumento.

El instrumento para evaluar proyectos de investigación experimental educativa, está constituido por seis grandes títulos: el problema, relación problema-diseño, presuposiciones, extracción de la muestra, instrumentos y técnicas de observación y estadística.

El primer gran título es el problema y como su nombre lo indica contiene las pautas de evaluación relacionadas con el problema propiamente dicho, estas pautas son la significación del problema por una parte y por otra la definición del problema. Cada una de estas dos pautas se encuentran a su vez divididas en sub-pautas que son consideradas como puntos importantes dentro de cada pauta a evaluar. Para la pauta significación del problema las sub-pautas correspondientes son: el propósito y el origen del problema; por otra parte para la pauta definición del problema las sub-pautas correspondientes son: la pregunta, la hipótesis, las variables dependientes o independientes. Junto a cada sub-pauta (ver Anexo A) están dados dos adjetivos bipolares, es decir adjetivos cuyo significado son extremos en cuanto a una cualidad considerada, que permitan al evaluador fijar su atención sobre los mismos aspectos pertinentes de cada sub-pauta.

El segundo gran título es la relación problema-diseño que contiene la pauta de evaluación que corresponde al grado

de adecuación del diseño con respecto al problema que se investiga. Esta pauta contiene las sub-pautas: hipótesis, procedimiento, tratamiento y duración. Junto a cada sub-pauta (ver Anexo A) están dados los adjetivos bipolares apropiados.

El tercer gran título presuposiciones, contiene la pauta de evaluación que corresponde al grado en que el diseño controla las variables. Está compuesta de una serie de sub-pautas que son: variables independientes, aplicación de tratamiento, factores de grupo o profesor, prejuicio del sujeto o experimentador, factores extraños y factores individuales. Junto a cada sub-pauta (ver Anexo A) están dados los adjetivos bipolares apropiados.

El cuarto gran título extracción de la muestra, contiene la pauta de evaluación que permite evaluar el grado de adecuación de la muestra seleccionada con respecto al diseño y propósito de la investigación. Contiene las sub-pautas: población, extracción de la muestra, aplicación de tratamiento, tamaño y características. Todas ellas con sus respectivos adjetivos bipolares (ver Anexo A).

El quinto gran título instrumentos y técnicas de observación, permite evaluar el grado de validez y confiabilidad de las técnicas de observación y de los instrumentos. Las sub-pautas que incluye son: validez y confiabilidad para la población, procedimientos de recolección de datos. Junto a cada sub-pauta (ver Anexo A) están dados los adjetivos bipolares apropiados.

Finalmente el sexto gran título estadística, contiene la pauta de evaluación que permite evaluar el grado en el que los estadísticos empleados para analizar los datos se ajustan a las características de las variables. Contiene las sub-pautas: tests estadísticos y datos. Cada sub-pauta está acompañada de sus respectivos adjetivos bipolares (ver Anexo A).

Cada pauta de evaluación contiene una escala de cinco puntos dónde el evaluador hará su evaluación. El punto uno de la escala corresponde a excelente, el punto dos a muy bueno, el punto tres a bueno, el punto cuatro a regular y el punto cinco a malo. El puntaje total de la evaluación estará dado por la suma de los puntajes obtenidos sobre cada una de las escalas de cinco puntos de las siete pautas de evaluación dando un puntaje de cinco al punto uno de la escala, un puntaje de cuatro al punto dos de la escala, un puntaje de tres al punto tres de la escala, un puntaje de dos al punto cuatro de la escala y un puntaje de uno al punto cinco de la escala. El puntaje total variará de un mínimo de uno a 35 puntos. Estos puntajes fueron reducidos a una escala de 0-100 para poder compararlos con los puntajes dados por los evaluadores que no utilizaron el instrumento.

Como parte del instrumento propuesto para evaluar proyectos de investigación experimental educativa, el investigador hizo una guía para el instrumento adaptado de Suydam (ver Anexo B). Esta guía intenta dar los aspectos fundamentales de cada una de las pautas de evaluación consideradas. También ayuda a

que se constituya en una primera aproximación global de la investigación por parte del estudiante que inicia una investigación y le permita situarse rápidamente en cada uno de los aspectos que constituirán su investigación.

Inventarios--con el objeto de dar un primer paso en la evaluación de los conocimientos de los cursos de Medición, Evaluación e Investigación Educativa que permitieran caracterizar mejor a la población bajo estudio, se aplicaron dos inventarios denominados uno Inventario de medición del conocimiento de vocabulario básico en medición y evaluación educativa y otro Inventario de medición del conocimiento de vocabulario básico en investigación educativa. Para facilidad de ubicación del lector se reescribe la premisa que fundamenta la aplicación de dichos inventarios y que fue dada en el diseño experimental. La premisa es: el conocimiento del vocabulario ha sido frecuentemente usado como un índice del conocimiento de la persona de un campo especializado Ebel (1960).

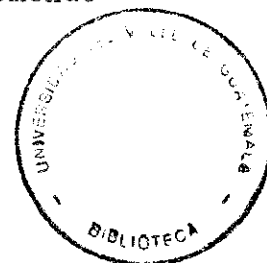
El Inventario de medición del conocimiento de vocabulario básico de medición y evaluación educativa Ebel (1960)-- este inventario está constituido de 149 términos de uso común en medición educativa, tomados de varios glosarios de términos de medición, prevaleciendo para su selección el juicio del constructor. El investigador para facilitar el uso del inventario agregó una lista complementaria de los 149 términos en español. Junto a cada término del vocabulario se encuentra una escala de cuatro

puntos cada uno de los cuales en grado ascendente indicó el conocimiento del término considerado por parte del sujeto de la población bajo estudio (ver Anexo C).

El inventario de medición del conocimiento de vocabulario básico en investigación educativa--está constituido de 120 términos de uso común en investigación educativa, tomados de varios glosarios de términos de investigación existentes en libros de investigación educativa, prevaleciendo para su selección en el juicio del investigador la experiencia en el curso de Medición, Evaluación e Investigación Educativa desarrollada principalmente en la materia de Diseño experimental. Junto a cada término del vocabulario se encuentra una escala de cuatro puntos cada uno de los cuales en grado ascendente indicó el conocimiento del término considerado por parte del sujeto de la población bajo estudio (ver Anexo D).

Recolección de datos

A cada evaluador, le fue entregada una copia del instrumento de evaluación de proyectos de investigación experimental para emplearla en la evaluación de cada artículo. A los evaluadores que aplicaron el instrumento sólo se les hicieron algunas observaciones acerca del uso del instrumento, pero no les fue dado un entrenamiento previo del instrumento y a los evaluadores que no aplicaron el instrumento se les recomendó que calificaran cada artículo en la escala de 0-100.



El grupo experimental que empleó el instrumento, así como el de control, que no lo empleó, efectuaron su evaluación independientemente tanto de un grupo a otro como entre cada evaluador y los demás. El tiempo empleado por cada evaluador fue a discreción para que se realizara la evaluación totalmente.

Análisis de los datos

En este subtítulo se presenta la relación formal entre las hipótesis planteadas y los resultados obtenidos, los resultados del test F de diferencias de desviaciones estándar y en último término el resumen de los resultados obtenidos de la aplicación de los inventarios descritos en el sub-título de instrumentos.

Para facilidad en la lectura del presente informe de investigación, los enunciados de las hipótesis planteadas serán escritas nuevamente antes de dar los resultados que permitan establecer la relación formal entre las hipótesis y sus resultados.

Hipótesis uno: la media de los puntajes obtenidos por los evaluadores que aplicaron el instrumento de evaluación de proyectos de investigación experimental es igual a la media de los puntajes obtenidos por los evaluadores que no aplicaron el instrumento.

Los resultados obtenidos en esta investigación se encuentran tabulados en el cuadro 4.1.

Las medias se testaron con el modelo "t" de Student mediante un ensayo bilateral.

CUADRO 4.1

Medias de la variable independiente
"t" observado y "t" crítico

	Variable independiente		"t" observado	"t" crítico
	Sin instrumento Media	Con instrumento Media		
Artículo I ^{a/}	73.33	86.67	3.12 ⁺	2.78
Artículo II ^{b/}	83.33	80.00	.47	2.78
Artículo III ^{c/}	88.33	88.57	.06	2.78

+ Significativo al nivel del 0.05

a/ El "dibujo de la persona" en esquizofrénicos y normales.

b/ Influencia del prejuicio del experimentador sobre la observación de un evento físico.

c/ El MMPI (español) aplicado a jóvenes mexicanos: influencias de sexo, edad y nivel de inteligencia.

Los resultados encontrados, en el cuadro 4.1, al testar la hipótesis uno bajo estudio para el artículo I--el "dibujo de la persona" en esquizofrénicos y normales-- fue que ésta resultó significativa; para el artículo II--influencia del prejuicio del experimentador sobre la observación de un evento físico-- fue que la hipótesis resultó no significativa y para

el artículo III--el MMPI (español) aplicado a jóvenes mexicanos: influencia de sexo, edad y nivel de inteligencia-- la hipótesis resultó no significativa.

Hipótesis dos: la desviación estándar de los puntajes obtenidos por los evaluadores que aplicaron el instrumento de evaluación de proyectos de investigación experimental es igual a la desviación estándar de los puntajes obtenidos por los evaluadores que no aplicaron el instrumento.

En el cuadro 4.2 se presentan los resultados obtenidos en el test F de significación de diferencia de desviaciones estándar para muestras pequeñas.

Las desviaciones estándar se testaron con el estadístico F mediante un ensayo bilateral, a un nivel de significación de 0.05.

CUADRO 4.2

Test F de significación de diferencia de desviaciones estándar para muestras pequeñas

Variable independiente						
Sin instrumento	Con instrumento			F observada	F crítico	
Media D.E. n = 9	Media D.E. n = 9		g.l.			
81.66	7.44	85.08	11.1	8	2.24	3.23

El resultado encontrado, en el cuadro 4.2, al testar la significación de la diferencia de desviaciones estándar para los puntajes obtenidos por los dos tipos de evaluadores no es significativa al nivel de 0.05, es decir, no hay prueba alguna de que los dos grupos difieran en variabilidad.

Resumen de resultados de inventarios

Con el objeto de calcular los porcentajes de conocimiento de cada término, de los inventarios de medición del conocimiento de vocabulario básico en medición y evaluación educativa y de medición del conocimiento de vocabulario básico en investigación educativa, por parte de los 16 alumnos de la maestría se asignó tres puntos para cada término que podía ser identificado y un punto para cada término que únicamente haya sido alguna vez escuchado o visto. Combinando los puntajes de los 16 alumnos en cada ítem y dividiendo por el máximo puntaje posible, se obtiene la figura de conocimiento de cada término y que es dada en los cuadros 4.3 y 4.4 para los respectivos inventarios ya mencionados.

CUADRO 4.3

Figura de conocimiento para el inventario de medición
del conocimiento de vocabulario básico en
medición y evaluación educativa

Vocabulario	Figura de conocimiento %
1. Análisis factorial	97.6%
2. Aprendizaje de memoria	97.6
3. Atenuación	40.5
4. Batería de aptitudes diferenciales	95.2
5. Batería de tests	100.0
6. Calificación ponderada	90.5
7. Clasificación Q (una escala)	69.0
8. Clave empírica	64.3
9. Clave para calificar	100.0
10. Cociente de inteligencia IQ	97.6
11. Cociente de rendimiento	85.7
12. Coeficiente de confiabilidad	100.0
13. Coeficiente de correlación	100.0
14. Coeficiente ϕ (phi)	69.0
15. Contador gráfico de ítems	52.4
16. Confiabilidad de división por mitad	100.0
17. Confiabilidad test-retest	100.0
18. Conjunto de respuestas	83.3
19. Consistencia interna	92.9

Continúa Cuadro 4.3

Vocabulario	Figura de conocimiento %
20. Corrección por adivinación	88.1
21. Correlación ítem-test	83.3
22. Criterio	95.2
23. Criterio interno	83.3
24. Conteo de respuestas	88.1
25. Cuestionario	100.0
26. Cuartiles	100.0
27. Crédito por medio de examen	85.7
28. Deciles	95.2
29. Desviación	100.0
30. Desviación estándar	100.0
31. Destrezas básicas	88.1
32. Diagrama de dispersión	95.2
33. Determinantes específicos	57.1
34. Diferencia significativa	97.6
35. Dispersión	100.0
36. Distractores	100.0
37. Distribución asimétrica	100.0
38. Distribución bimodal	100.0
39. Distribución de frecuencia	100.0
40. Distribución normal	100.0
41. Distribución de puntajes	100.0

Continúa Cuadro 4.3

Vocabulario	Figura conocimiento %
42. Distribución rectangular	52.4
43. Edad mental	95.2
44. Efecto de Halo	71.4
45. Efecto de práctica	83.3
46. Ejercicio de clasificación	83.3
47. Ejercicio de pareamiento	88.1
48. Ejercicio de prueba	85.7
49. Enunciado del ítem	100.0
50. Error de muestreo	97.6
51. Error estándar de medición	95.2
52. Error de medición	95.2
53. Error probable	81.0
54. Error de varianza	95.2
55. Escala	95.2
56. Escala de edad	88.1
57. Escala de calidad	66.7
58. Escala gráfica de clasificación	78.6
59. Estadístico	100.0
60. Evaluación	100.0
61. Evaluación subjetiva	100.0
62. Examen	100.0
63. Examen de ensayo	100.0

Continúa Cuadro 4.3.

Vocabulario		Figura de conocimiento %
64.	Examen externo	81.0
65.	Escala de puntos	92.9
66.	Formas equivalentes	100.0
67.	Fórmula de Kuder-Richardson	97.6
68.	Formula para calificar (cuando se restan % malas de buenas)	97.6
69.	Fórmula de Spearman-Brown	97.6
70.	Frecuencia acumulada	97.6
71.	Grupo de control	100.0
72.	Heterogeneidad	100.0
73.	Homogeneidad	100.0
74.	Incidente crítico	59.5
75.	Indice de discriminación	97.6
76.	Inventario	95.2
77.	Indice de dificultad	97.6
78.	Item de falso-verdadero	100.0
79.	Item de mejor respuesta	92.9
80.	Item de selección múltiple	100.0
81.	Item de test	97.6
82.	Máquina calificadora IBM	97.6
83.	Marca	95.2
84.	Media	100.0
85.	Mediana	100.0

Continúa Cuadro 4.3

Vocabulario		Figura de conocimiento %
86.	Medición	100.0
87.	Moda	100.0
88.	Muestra aleatoria	100.0
89.	Muestra representativa	100.0
90.	Norma	100.0
91.	Norma de edad	100.0
92.	Norma de grado	97.6
93.	Números aleatorios	100.0
94.	Perfil	97.6
95.	Pertinencia	78.6
96.	Probabilidad	100.0
97.	Promedio	100.0
98.	Prueba de rapidez	100.0
99.	Puntaje	100.0
100.	Puntaje burdo	100.0
101.	Puntaje compuesto	88.1
102.	Puntaje comparable	97.6
103.	Puntaje crítico	85.7
104.	Puntaje derivado	78.6
105.	Puntaje estándar	97.6
106.	Puntaje estándar normalizado	97.6
107.	Puntaje estaninos	85.7

Continúa Cuadro 4.3

Vocabulario	Figura de conocimiento %
108. Puntaje de razón	69.0
109. Puntaje T	95.2
110. Puntaje verdadero	88.1
111. Puntajes z	73.8
112. Pureza (de un test)	92.9
113. Rango de puntajes	97.6
114. Rango percentil	97.6
115. Rasgo, característica	88.1
116. Regresión	97.6
117. Sigma	97.6
118. Sociometría	100.0
119. Sub-test	97.6
120. Sistema de honestidad	57.1
121. Tabla de especificaciones	100.0
122. Tabla de expectación	81.0
123. Test	100.0
124. Test de analogías	97.6
125. Test de aptitudes	100.0
126. Test de calificación simultánea	83.3
127. Test de completación	92.9
128. Test de ejecución (de actuación)	100.0
129. Test de interpretación	69.0

Continúa Cuadro 4.3

Vocabulario	Figura de conocimiento %
130. Test de libro abierto	100.0
131. Test de maestría	100.0
132. Test estandarizado	100.0
133. Test de tiempo límite	97.6
134. Test de diagnóstico	100.0
135. Test no verbal	97.6
136. Test Libre de cultura	76.2
137. Test de potencia	64.3
138. Test proyectivo	97.6
139. Test objetivo	100.0
140. Test de rendimiento	100.0
141. Test de Rorschach	92.9
142. Técnica de selección forzada	64.3
143. Test limitado por ejecución	69.0
144. Validación cruzada	69.0
145. Validez aparente	95.2
146. Validez de contenido	97.6
147. Validez estadística	90.5
148. Validez	90.5
149. Varianza	92.9

Combinando las respuestas de los estudiantes de la maestría se encontró que el 3.1 por ciento indicaron ningún

contacto previo del término, 3.1 por ciento solamente habían escuchado del término, 10.67 por ciento tuvieron algún conocimiento del término y 83.57 por ciento tuvieron un entendimiento cabal del término. La media para la figura de conocimiento de todos los ítemes fue aproximadamente de 90 por ciento.

CUADRO 4.4

Figura de conocimiento para el inventario de medición
del conocimiento de vocabulario básico en
investigación educativa

Vocabulario	Figura de conocimiento %
1. Análisis de contenido	76.2
2. Bloqueo incompleto	54.8
3. Comparación	100.0
4. Comparación ortogonal	100.0
5. Confiabilidad	100.0
6. Confiabilidad del observador	85.7
7. Constructo hipotético	85.7
8. Correlación	100.0
9. Cuadrados conjuntos	7.14
10. Cuadrado estándar	47.6
11. Cuasiexperimento	97.6
12. Cuestionario	100.0

Continúa Cuadro 4.4

Vocabulario	Figura conocimiento %
13. Datos básicos de referencia	78.6
14. Datos iniciales de referencia	73.8
15. Definición operacional	85.7
16. Desviación	97.6
17. Diseño de postest solamente	100.0
18. Diseño de grupo paralelo	95.2
19. Diseño factorial	97.6
20. Diseño pretest-postest	100.0
21. Distribución de muestras	97.6
22. Distribución normal multivariada	66.7
23. Eficiencia relativa de un estadístico	78.6
24. Encuesta	100.0
25. Entrevista	100.0
26. Entrevista estructurada	100.0
27. Entrevista no estructurada	100.0
28. Error	100.0
29. Error experimental	97.6
30. Error por azar	97.6
31. Error sistemático	78.6
32. Error tipo I	100.0
33. Error tipo II	100.0
34. Escala graduada	90.5

Continúa Cuadro 4.4

Vocabulario	Figura de conocimiento %
35. Estadística	100.0
36. Estimador	100.0
37. Estimador consistente	88.1
38. Estimador eficiente	83.3
39. Estimador insesgado	97.6
40. Evaluación por rangos	83.3
41. Fuente primaria	95.2
42. Fuente secundaria	95.2
43. Generalización	97.6
44. Grado de libertad	97.6
45. Grupo de control	100.0
46. Hipótesis	100.0
47. Hipótesis declarativa	78.6
48. Hipótesis de investigación	100.0
49. Hipótesis estadística	100.0
50. Hipótesis interrogativa	85.7
51. Hipótesis nula	100.0
52. Información relativa	71.4
53. Interacción	97.6
54. Intervalo de confianza	100.0
55. Investigación aplicada	100.0
56. Investigación básica	100.0

Continúa Cuadro 4.4

Vocabulario	Figura de conocimiento %
57. Investigación de campo	90.5
58. Investigación de laboratorio	92.9
59. Investigación descriptiva	100.0
60. Investigación experimental	100.0
61. Investigación fundamental	100.0
62. Investigación histórica	100.0
63. Investigación pura	81.0
64. Límites de confianza	100.0
65. Matriz de varianza-covarianza	76.2
66. Método de distribución libre	59.5
67. Medición	97.6
68. Método no paramétrico	92.9
69. Modelo de efectos aleatorios	76.2
70. Modelo de efectos fijos	73.8
71. Modelo estadístico	100.0
72. Modelo mixto	71.4
73. Muestra	100.0
74. Muestra aleatoria	100.0
75. Nivel de significancia	97.6
76. Observación	100.0
77. Parámetro	100.0
78. Pares combinados o apareamiento	90.5
79. Placebo	73.8

Continúa Cuadro 4.4

Vocabulario	Figura de conocimiento %
80. Población	100.0
81. Población hipotética	92.9
82. Postulado	92.9
83. Pregunta abierta	97.6
84. Pregunta cerrada	97.6
85. Procedimiento interferente	69.0
86. Probabilidad	100.0
87. Prueba estadística	100.0
88. Región crítica	100.0
89. Región de aceptación	100.0
90. Relación de causa-efecto	97.6
91. Relación funcional	69.0
92. Replicación	83.3
93. Selección al azar	100.0
94. Significación estadística	97.6
95. Significación práctica	71.4
96. Suma de cuadrados ajustada	100.0
97. Teoría	100.0
98. Teoría de la decisión estadística	83.3
99. Test de potencia	69.0
100. Test estadístico	100.0
101. Test de razón verosimilitud	100.0

Continúa Cuadro 4.4

Vocabulario	Figura de conocimiento %
102. Test más poderoso	88.1
103. Tratamiento	97.6
104. Tratamientos cruzados	76.2
105. Tratamiento incluido	66.7
106. Unidad de análisis	69.0
107. Unidad de muestreo	88.1
108. Universo	100.0
109. Validez	100.0
110. Validez externa	97.6
111. Validez interna	97.6
112. Valor crítico	97.6
113. Valor esperado	100.0
114. Variable	100.0
115. Variable-estímulo	85.7
116. Variable independiente	100.0
117. Variable aleatoria	97.6
118. Variable dependiente	100.0
119. Variable interviniente	90.5
120. Variable respuesta	66.7

Combinando las respuestas de los estudiantes de la maestría se encontro que el 2.7 por ciento indicaron ningún contacto

previo del término, 3.6 por ciento solamente habían escuchado del término, 14.9 por ciento tuvieron algún conocimiento del término y 80.1 por ciento tuvieron un entendimiento cabal del término. La media para la figura de conocimiento de todos los ítemes fue aproximadamente de 90 por ciento.

CAPÍTULO V

DISCUSIÓN, CONCLUSIONES, LIMITACIONES Y RECOMENDACIONES

En el capítulo anterior se concluyó parcialmente este informe hasta el análisis de datos. En este capítulo se presenta la discusión de los resultados encontrados en el estudio, se obtienen además algunas conclusiones, se señalan ciertas limitaciones del estudio y se hace una serie de recomendaciones para otros estudios relacionados con el tema de evaluación de proyectos de investigación educativa.

Discusión

El hecho de que la hipótesis uno haya resultado significativa para el artículo I indica que el uso o no del instrumento parece establecer una diferencia real en la evaluación del artículo I. En cambio para los artículos II y III la hipótesis uno no es rechazada, es decir, que no existe evidencia para indicar una diferencia real en la evaluación de informes cuando ésta se hace con o sin instrumento. Por otra parte, el hecho de que el test F de significación de diferencia de desviaciones estándar con que se testó la hipótesis dos no haya resultado significativo permite pensar que no existió beneficio en el uso del instrumento propuesto en cuanto a la variabilidad de los puntajes obtenidos por los evaluadores que hicieron uso de él, con respecto a los evaluadores que no lo usaron.

Una de las causas que talvez influyó en estos resultados es el hecho de que el presente trabajo es una primera investigación del investigador y está hecha bajo las limitaciones de experiencia, tiempo y circunstancias favorables propias de este tipo de trabajo.

Las causas que talvez influyeron en estos resultados se pueden deber a las restricciones que tuvo el investigador en la selección de literatura científica pertinente al estudio. Una de ellas podría ser la escasez de literatura científica en español de la biblioteca de la Universidad del Valle de Guatemala.

Otra causa posible de los resultados obtenidos fue la falta de dominio del inglés de la población bajo estudio que obligó al investigador a seleccionar los artículos sólo de dos fuentes, con el agravante de que dichas fuentes eran de psicología, por lo que posiblemente no existió una verdadera representatividad del contenido de los artículos en cuanto al área educativa.

Un aspecto más que pudo influir en los resultados obtenidos es el hecho de que la muestra seleccionada, haya sido pequeña y que dentro de ella hayan quedado incluidos dos psicólogos quienes por razón de su profesión tuvieron por lo tanto mayor familiaridad con el material evaluado, este hecho pudo haber dado un sesgo al estudio.

Otra causa factible del sesgamiento de los resultados es el posible contacto previo de los evaluadores con algunos artículos a evaluar debido a la revisión de literatura llevada por el grupo de maestría, de donde se extrajo la muestra de evaluadores, para la elaboración de su tesis de grado.

Las condiciones en que se hizo el estudio pudieron haber sido otro factor que haya influido en los resultados obtenidos; las condiciones de evaluación se dieron en un período de trabajos y exámenes para la población bajo estudio.

La falta de entrenamiento con el instrumento probablemente haya contribuido a los resultados obtenidos.

Otras variables a considerar como factores de sesgo del estudio fueron los antecedentes académicos, experiencia como evaluadores, de los integrantes de la población bajo estudio.

La carencia de proyectos a evaluar hizo necesario tomar como elementos de evaluación informes de investigación experimental que necesariamente introdujeron cierta perturbación a los evaluadores por encontrar otras características en el artículo a evaluar además de las propias del proyecto.

Conclusiones

Basado en la evidencia estadística que fue obtenida al testar las hipótesis y analizar los datos, se puede concluir:

1. El instrumento propuesto no permitió obtener una diferencia significativa entre los puntajes

obtenidos por los evaluadores que utilizaron instrumento y los que no lo emplearon.

2. El instrumento propuesto no permitió obtener una menor dispersión de los puntajes obtenidos por los evaluadores que usaron el instrumento y los que no lo emplearon.

Limitaciones del estudio

La realización del presente estudio se hizo bajo cierto número de limitaciones que enmarcan su generalización:

1. El estudio sólo se hizo sobre una población perteneciente al grupo de maestría de la Universidad del Valle de Guatemala por lo que sus conclusiones deben quedarse dentro de esta población.
2. El instrumento propuesto para evaluar proyectos de investigación experimental educativa dadas sus características propias no permite que los resultados obtenidos sean aplicables a otro tipo de instrumento de evaluación que no tenga propósitos y características semejantes al propuesto.

Recomendaciones

Las siguientes recomendaciones propuestas por el investigador, para futuras investigaciones en el tema de evaluación de proyectos de investigación, toman la siguiente forma:

1. En la selección de artículos a evaluar se podría pensar en traducciones de alta calidad, de aquellos artículos que fueran seleccionados al azar de entre toda la literatura existente en la Biblioteca de la Universidad del Valle de Guatemala.
 2. La carencia de proyectos para ser evaluados puede ser superada con la creación de un archivo de proyectos de investigación que se podría constituir con los proyectos de investigación presentados por los alumnos de la maestría, tanto en el curso de Diseño experimental como en el Seminario.
 3. Se podría ampliar el origen de los evaluadores considerando como posibles participantes del experimento al cuerpo de investigadores de las diferentes Universidades de Guatemala que se encuentren involucrados en la evaluación de investigación.
 4. Con el objeto de superar la falta de entrenamiento con el instrumento propuesto por parte de los evaluadores se sugiere la inclusión dentro del curso de Diseño experimental de unas sesiones dedicadas al empleo de él y a la crítica del mismo para lograr su buen uso.
 5. Posiblemente se aumentaría la riqueza de la información con la inclusión de una escala de cinco
-

puntos entre los adjetivos bipolares de cada sub-pauta. Esta escala permitiría al evaluador graduar su criterio.

6. La realización de estudios piloto para mejorar la confiabilidad y validez del instrumento.

ANEXOS

ANEXO A

INSTRUMENTO PARA EVALUAR PROYECTOS DE INVESTIGACIÓN

INSTRUCCIONES

El siguiente instrumento incluye siete pautas de evaluación. Antes de evaluar cada pauta considere cada uno de los incisos que incluye ésta, seguidos de adjetivos que indican los extremos del continuo sobre el cual debe evaluar dichos aspectos. No responda a estos incisos, sólo están destinados a enfocar la atención de todos los evaluadores sobre los mismos aspectos pertinentes de cada pauta. La calidad del proyecto se evaluaría en términos de cada una de las siete pautas de evaluación usando para cada una de ellas la siguiente escala de cinco puntos:

1. Excelente: incluye todos los requisitos de la pauta de evaluación planteada; nada esencial se podría agregar.
2. Muy bueno: incluye la mayoría de los requisitos.
3. Bueno : incluye aproximadamente la mitad de los requisitos.
4. Regular : contiene muy pocos requisitos.
5. Malo : no incluye ninguno o demasiados pocos de los requisitos.

Para indicar su evaluación de las pautas en las páginas siguientes, tache con una cruz el número seleccionado, en la escala de cinco puntos, su evaluación para cada punto, por lo que son siete evaluaciones para cada proyecto.

I. EL PROBLEMA

1. Grado en el que el problema seleccionado es significativo para la teoría o para la práctica.

1 - 2 - 3 - 4 - 5

a. Propósito (importante - no importante)

b. Origen del problema

1) Justificación (válida - no válida)

2) Investigación previa (adecuada - inadecuada)

2. Claridad con la que se define el problema

1 - 2 - 3 - 4 - 5

a. Pregunta (operacional - vaga)

b. Hipótesis (relevante - irrelevante)

(lógica - ilógica)

c. Variable(s) independiente(s) (operacional - vaga)

d. Variable(s) dependiente(s) (operacional - vaga)

II. RELACIÓN PROBLEMA - DISEÑO

3. Grado en el que el diseño es adecuado para investigar el problema planteado.

1 - 2 - 3 - 4 - 5

- a. Hipótesis (testable - no testable)
- b. Procedimiento (claro - no claro)
- c. Tratamiento (replicable - no replicable)
- (adecuado - inadecuado)
- d. Duración (adecuada - inadecuada)

III. PRESUPOSICIONES

- 4. Grado en el que el diseño controla las variables

1 - 2 - 3 - 4 - 5

- a. Variable(s) independiente(s) (contaminada - no contaminada)
- b. Aplicación de tratamiento (riguroso - no riguroso)
- c. Factores de grupo o profesor (controlado - no controlado)
- d. Prejuicios del sujeto o experimentador (controlado - no controlado)
- e. Factores extraños (controlado - no controlado)
- f. Factores individuales (controlado - no controlado)

IV. EXTRACCIÓN DE LA MUESTRA

- 5. Grado en que la muestra seleccionada es adecuada para el diseño y los propósitos de la investigación.

1 - 2 - 3 - 4 - 5

- a. Población (adecuada - inadecuada)
- b. Extracción de la muestra (azar - no al azar)
- c. Aplicación de tratamiento (azar - no al azar)
- d. Tamaño (adecuada - inadecuada)
- e. Características (adecuada - inadecuada)

V. INSTRUMENTOS Y TÉCNICAS DE OBSERVACIÓN

- 6. Validez y confiabilidad de las técnicas de observación y de los instrumentos.

1 - 2 - 3 - 4 - 5

- a. Instrumento o técnica
 - 1) Descripción (excelente - mala)
 - 2) Validez (adecuada - inadecuada)
 - 3) Confiabilidad para la población (excelente - deficiente)
- b. Procedimiento de recolección de datos (controlado - sin control)

VI. ESTADÍSTICA

- 7. Grado en el que los estadísticos empleados para analizar los datos se ajustan a las características de las variables.

1 - 2 - 3 - 4 - 5

a. Tests estadísticos

- 1) Presuposiciones básicas . (satisfechas - no satisfechas)
- 2) Relación con el diseño (adecuado - inadecuado)

b. Datos

- 1) Nivel de significación (especificado - no especificado)
(adecuado - inadecuado)

ANEXO B

GUÍA DEL INSTRUMENTO PARA EVALUAR PROYECTOS DE INVESTIGACIÓN

Esta guía está organizada en rubros y subrubros presentados en el mismo orden en que aparecen en el instrumento de proyectos de investigación.

I. EL PROBLEMA

1. Grado en el que el problema seleccionado es significativo para la teoría o para la práctica.

a. Propósito: el problema surge de una situación en la que se reconoce que existen dificultades por resolver y que debe ser digna de estudio. La importancia del problema se puede medir por su relación con las necesidades de la institución, sistema, de la región, del país o de una determinada rama del conocimiento.

b. Origen del problema

1) Justificación: el interés personal del investigador y/o su deseo de servir a la sociedad pueden ser fuente de problemas a investigar. En uno y otro caso los estudios anteriores darán la pauta de validez de dichos problemas seleccionados por el investigador.

2) Investigación previa: debe proporcionar evidencia de búsqueda de literatura relevante y pertinente al problema estudiado; no tiene que ser una revisión completa de toda la información pedagógica disponible y de las publicaciones que tienen relación con los factores del problema pero si debe incluir el investigador en este rubro lo que se conoce de importante sobre la materia. Este resumen de estudios previos se presenta en forma hilvanada que sugiera un marco conceptual más bien que en orden cronológico. A la información más reciente se le debe dar la mayor atención debido a que presumiblemente proporciona los más recientes descubrimientos e incluye las fuentes más antiguas; no se hace necesaria una larga lista de referencias.

2. Claridad con la que se define el problema.

a. Pregunta: el investigador debe aprender a reconocer y definir el problema, puesto que, para llevar a cabo una investigación es necesario identificar con exactitud la naturaleza y dimensiones del problema. Definir un problema significa especificarlo en detalle y con precisión, ésto servirá para complementar los juicios sobre su valor y practicabilidad. Un planteamiento apropiado del problema, en forma declarativa o interrogativa es probablemente la parte más importante

del proyecto de investigación. Debe por lo tanto estar en términos específicos, precisos y concisos de tal forma que conduzca naturalmente a todos los pasos subsecuentes. La pregunta crucial que destaca entre todas las que el estudio debe responder, la que en realidad incluye a todas las demás, es el problema considerado.

b. Hipótesis: las hipótesis son posibles soluciones del problema que se expresan como generalizaciones o proposiciones; ciertos elementos o relaciones de las hipótesis son hechos observados, en tanto que otros son conceptuales. La cantidad y la calidad de las hipótesis que se formulan dependerán del alcance y riqueza de los conocimientos que el investigador posea y pueda aplicar al análisis del problema, la versatilidad y discriminación con que se seleccionan los hechos; se elaboren los conceptos; se identifiquen relaciones; y se las organice según pautas de explicación significativas. Las hipótesis deben formularse en forma afirmativa, clara y precisa; una hipótesis posee mayor valor si establece una generalización que pueda aplicarse en muchas áreas de la educación. En general una buena hipótesis tiene algunas características básicas: debe ser razonable, de acuerdo con los hechos conocidos, enunciada de modo que se pueda someter a prueba y enunciada en términos tan sencillos como sea posible.

- c. Variable(s) independiente(s): un problema debe referirse a relaciones entre dos o más variables. En el sentido más general una variable es algo que puede cambiar cuantitativa o cualitativamente. Este algo pueden ser atributos, características o propiedades que pueden cambiar en cada individuo o entre individuos. Por su uso, las variables pueden ser independientes, dependientes o intervinientes. La variable independiente (predictora) en los estudios experimentales es la sometida a variación.
- d. Variable(s) dependiente(s): es el atributo, propiedad o característica que se trata de cambiar mediante la manipulación de la variable independiente. Se llama también criterio porque es el fenómeno que interesa en última instancia al investigador. Esta variable está dada por los puntajes recolectados por medio del instrumento aplicado.

II. RELACIÓN PROBLEMA DISEÑO

3. Grado en el que el diseño es adecuado para investigar el problema planteado.
- a. Hipótesis: si la hipótesis experimental incluye variables que se pueden medir, se puede someter a prueba. Una razón para pensar que la hipótesis no se puede someter a prueba es la posibilidad de que una o

más variables que incluya la hipótesis no puedan ser medidas por no contar los investigadores con el o los instrumentos adecuados. La característica de una hipótesis adecuada es que pueda ser sometida a prueba. Por medio de la investigación, la hipótesis experimental no se puede confirmar, sólo se puede rechazar o dejar de rechazar.

b. Procedimiento: en esta sección se describe en detalle la forma como se procederá para someter a prueba la hipótesis. A veces hay que adoptar métodos que implican presuposiciones especiales y éstas se deben formular explícitamente; es muy importante que se haga una selección adecuada de los estadísticos de prueba. Una cuestión experimental no se puede contestar por medio de una descripción. El rubro "procedimiento" es el lugar adecuado para entrar en detalles de las técnicas a adoptar y para describir los instrumentos a emplear o a constituirse para la recolección de datos; para dar las fundamentaciones de confiabilidad y validez de los instrumentos. Los pasos seguidos para desarrollar el proyecto se deben describir en suficiente detalle para que puedan ser completamente entendidos. En la medida que ésta condición se dé, el experimento podrá ser repetido por otro investigador. No deben omitirse detalles importantes. El rubro procedimiento debe incluir una descripción completa de la forma como

se van a combinar en la ejecución del experimento: variables; sujetos; aparatos; instrumentos; locales; y experimentadores.

e. Tratamiento: recibe este nombre la forma en que se controlan ciertos aspectos de una situación, es decir, la manipulación de las variables independientes que se aplican a los sujetos. El tratamiento se refiere entonces a las condiciones a las cuales se somete al grupo experimental de sujetos. Investigar los efectos de los tratamientos sobre alguna forma específica de conducta de los sujetos es fundamentalmente el propósito del experimento.

d. Duración: el problema debe ser de tal naturaleza que pueda ser estudiado en un tiempo señalado para la entrega del proyecto. Se debe incluir en el proyecto una distribución de tiempo para la realización de las etapas del mismo.

III. PRESUPOSICIONES

4. Grado en el que el diseño controla las variables.

Es imposible para un investigador controlar todos los elementos en su proyecto, debe por lo tanto basar su trabajo sobre muchas suposiciones; es decir, suposiciones o cosas que se dan por admitidas con respecto a las diversas fases del procedimiento empleado en el estudio del

problema. Entre más rigurosa sea la investigación, menor será el número de suposiciones necesarias y cada una de ellas debe ser justificada. Algunas presuposiciones pueden permanecer implícitas, pero otras tales como la validez del instrumento, la adecuación de la muestra y el control de las variables, deben ser explícitas. Las investigaciones y/o los estudios experimentales que se basan en supuestos falsos no pueden producir resultados válidos. La necesidad en el diseño experimental de realizar una observación controlada conduce a la identificación y control de la variable independiente (predictora) y otras variables que influyen sobre la variable dependiente. El control establecido permite impedir que otro factor que no sea la variable independiente influya sobre la variable dependiente (predicha); permite también modificar la magnitud de la variable experimental y controlar las variables para describir cuantitativamente el alcance de su manifestación y sus efectos de interacción. El control se puede lograr dosificando la misma variable independiente para diversos grupos o controlando las variables no experimentales que puedan influir sobre la variable dependiente, manteniendo constantes las condiciones para el grupo experimental y de control o bien usando técnicas estadísticas como el método de correlación parcial y el análisis de varianza que permiten efectuar el análisis de situaciones en las que intervienen muchas variables. Una buena guía para analizar las

diferentes alternativas de diseño experimental en lo que respecta a los problemas de control de las variables externas y amenazas de validez lo constituye Donald Campbell y Julian Stanley, Diseños experimentales y cuasiexperimentales en la investigación social. Buenos Aires: Amorrortu editores, 1970.

IV. EXTRACCIÓN DE LA MUESTRA

5. Grado en que la muestra seleccionada es adecuada para el diseño y los propósitos de la investigación.
 - a. Población: para definir una población, debe ser claramente descrita en cuanto a sus características y tamaño; también deben ser explicadas las razones por las cuales se ha seleccionado la población bajo estudio. Una vez identificada la población con claridad, el investigador obtiene o elabora una lista completa, exacta y actualizada de todas las unidades que componen la población.
 - b. Extracción de la muestra: una vez identificada con claridad la población, en base a sus unidades se debe extraer una muestra, seleccionando un cierto número de unidades de dicha población. En el muestreo al azar, cuanto mayor sea el grado de heterogeneidad de las unidades y menor el tamaño de la muestra, habrá más posibilidades de que ésta resulte inadecuada. Algunas
-

investigadores describen como una muestra lo que son realmente subpoblaciones, aún cuando los denominen grupos de control y experimental; los descubrimientos de sus estudios puede decirse que se pueden generalizar a sus particulares subpoblaciones, y a todo aquello que tenga sus mismas características. El grupo de control se debe seleccionar con sumo cuidado.

d. Tamaño: el tamaño de la muestra, en general, dependerá de la variabilidad del fenómeno o característica bajo estudio, correspondiente a la población, del diseño experimental y el grado de predicción que se desee alcanzar.

e. Características: los miembros de la muestra deben ser elegidos al azar o sea que todos los integrantes de la población han de tener igual oportunidad de ser incluidos en la muestra; debido a que en la práctica esta selección aleatoria no se da en general, entonces la muestra seleccionada deberá ser al menos representativa de la población que se estudia.

V. INSTRUMENTOS Y RECOLECCIÓN DE DATOS

6. Validez y confiabilidad de las técnicas de observación y de los instrumentos.

a. Instrumentos o técnica

1) Descripción: una tarea importante del investigador

es seleccionar un instrumento de medición, cuya confiabilidad haya sido verificada, con el propósito de cuantificar las variables que ha seleccionado para su estudio. La selección del método de investigación influye en los procedimientos para medir las variables. Cualquier instrumento original se debe describir totalmente e ilustrar con ejemplos.

2) Validez: una instrumentación deficiente, es una falla seria pero común en muchas investigaciones educacionales. Para verificar que los instrumentos miden lo que se proponen medir tienen que ser sometidos a prueba en un sentido piloto o bien calibrados en alguna forma. Un matiz supuesto de validez no es suficiente. Varios instrumentos de validez deficiente usados en el mismo proyecto, no son necesariamente más válidos que uno, con validez verificada experimentalmente. Instrumentos prestados o modificados, elaborados por otro investigador, son buenos, mucho mejor aún si son instrumentos cuya validez se conoce.

3) Confiabilidad para la población: para verificar que los instrumentos medirán persistentemente, tienen que haber sido sometidos a prueba en un estudio piloto o bien calibrados en alguna forma. Instrumentos prestados o modificados elaborados por

otro investigador, son buenos, mucho mejor aún si son instrumentos cuya confiabilidad se conoce.

- b. Procedimiento de recolección de datos: se deben adoptar precauciones para garantizar que los datos hayan sido recogidos y registrados con objetividad y precisión. Estas precauciones están directamente motivadas por el procedimiento de recolección de datos seleccionado por el investigador.

VI. ESTADÍSTICA

7. Grado en que los estadísticos empleados para analizar los datos se ajustan a las características de las variables.

- a. Tests estadísticos

- 1) Presuposiciones básicas: las pruebas estadísticas a las que se someten las hipótesis requieren la consideración de ciertos supuestos que dependen del estadístico seleccionado por el investigador. Para el análisis de varianza se tienen supuestos como el de normalidad, de muestras aleatorias independientes, de desviaciones estándar iguales para las poblaciones; para la covarianza se trabaja con los supuestos de que la variable dependiente se distribuye en forma aproximadamente normal y que las varianzas de un grupo a otro son similares; en el

caso de los contrastes de medias con la "t" de Student se supone que las dos muestras son aleatorias y extraídas independientemente de una distribución normal y que existe igualdad de varianzas de las poblaciones de las que proceden las muestras (homogeneidad de varianzas).

2) Relación con el diseño: casi todas las investigaciones requieren algún tipo de análisis estadístico para evaluar la probabilidad de que la hipótesis experimental sea verdadera. En algunos casos para resolver un problema de carácter empírico, es preciso llevar a cabo un análisis bastante complejo; otras veces basta con efectuar un análisis muy simple y directo. La elección de uno u otro tipo de estadístico depende del problema que se plantea así como de la naturaleza de sus datos. Las técnicas más comúnmente usadas en estadística paramétrica son: el test "t", usado para encontrar si la diferencia entre dos medias muestrales es estadísticamente significativa. Existen dos tipos de estadísticos "t", el test-t para grupos no correlacionados, el cual es empleado para comparar medias muestrales cuando las muestras han sido extraídas independientemente de su población, y el test-t para grupos correlacionados el cual es empleado con dos muestras en las cuales los sujetos son pareados

o bien con dos mediciones repetidas obtenidas para los mismos sujetos.

El análisis de varianza se usa para comparar las medias de dos o más muestras y someter a prueba la hipótesis nula de que no existen diferencias estadísticamente significativas entre las medias obtenidas a partir de las muestras. El análisis multifactorial de varianza nos permite probar el efecto de más de una variable independiente y también el efecto de interacción de tales variables.

Entre los estadísticos inferenciales no paramétricos, el estadístico ji-cuadrada se emplea para encontrar las diferencias estadísticamente significativas entre proporciones de sujetos y/o eventos que caen dentro de diferentes categorías, comparando las frecuencias observadas y las frecuencias esperadas.

- b. Datos: en investigación educacional, el interés recae en estimar si los resultados se pueden deber al azar o no; y qué confianza se le puede atribuir a los datos recabados. La práctica corriente es permitir un máximo de 5 por ciento de azar, con el objeto de considerar que los resultados obtenidos no se deben al azar. Entre más pequeño sea el nivel de significancia considerado, la probabilidad de que el fenómeno observado se pueda atribuir al azar es menor.

ANEXO C

INVENTARIO DE MEDICIÓN DEL CONOCIMIENTO DE VOCABULARIO BÁSICO EN MEDICIÓN Y EVALUACIÓN EDUCATIVA

El siguiente inventario está constituido por 149 términos usados en medición y evaluación educativa y están dados en orden alfabético. Se da una escala de 4 puntos con el significado siguiente:

1. El término nunca lo ha encontrado en su experiencia en medición y evaluación.
2. Ha escuchado el término en cuestión pero no sabe lo que significa.
3. Puede identificar el término en general pero no explicar su significado específico.
4. Puede explicar su significado específico.

Tache su respuesta en la escala de 4 puntos.

VOCABULARIO

1. Análisis factorial	Factor analysis	1 - 2 - 3 - 4
2. Aprendizaje de memoria	Rote learning	1 - 2 - 3 - 4
3. Atenuación	Attenuation	1 - 2 - 3 - 4
4. Batería de aptitudes diferenciales	Differential apti- tudes battery	1 - 2 - 3 - 4
5. Batería de tests	Battery tests	1 - 2 - 3 - 4

6. Calificación ponderada	Weighted scoring	1 - 2 - 3 - 4
7. Clasificación Q (una escala)	Q sort	1 - 2 - 3 - 4
8. Clave empírica	Empirical key	1 - 2 - 3 - 4
9. Clave para la ca- lificar	Scoring key	1 - 2 - 3 - 4
10. Cociente de inteli- gencia IQ	IQ	1 - 2 - 3 - 4
11. Cociente de rendi- miento	Achievement quotient	1 - 2 - 3 - 4
12. Coeficiente de con- fiabilidad	Reliability coefficient	1 - 2 - 3 - 4
13. Coeficiente de co- rrelación	Correlation coefficient	1 - 2 - 3 - 4
14. Coeficiente ϕ (phi)	Phi coefficient	1 - 2 - 3 - 4
15. Contador gráfico de ítemes	Graphic item counter	1 - 2 - 3 - 4
16. Confiabilidad de di- visión por mitad	Split-halves reliability	1 - 2 - 3 - 4
17. Confiabilidad test-retest	Test-retest reliability	1 - 2 - 3 - 4
18. Conjunto de respues- tas	Response set	1 - 2 - 3 - 4
19. Consistencia interna	Internal consistency	1 - 2 - 3 - 4
20. Corrección por adi- vinación	Guessing correction	1 - 2 - 3 - 4
21. Correlación ítem-test	Item-test correlation	1 - 2 - 3 - 4
22. Criterio	Criterion	1 - 2 - 3 - 4
23. Criterio interno	Internal criterion	1 - 2 - 3 - 4
24. Conteo de respuestas	Response count	1 - 2 - 3 - 4
25. Cuestionario	Questionnaire	1 - 2 - 3 - 4
26. Cuartiles	Quartile	1 - 2 - 3 - 4

27. Crédito por medio de examen	Credit by examination	1 - 2 - 3 - 4
28. Deciles	Decile	1 - 2 - 3 - 4
29. Desviación	Deviation	1 - 2 - 3 - 4
30. Desviación estándar	Standar deviation	1 - 2 - 3 - 4
31. Destrezas básicas	Basic skills	1 - 2 - 3 - 4
32. Diagrama de dispersión	Dispersion diagram	1 - 2 - 3 - 4
33. Determinantes específicos	Specific determiners	1 - 2 - 3 - 4
34. Diferencia significativa	Significant difference	1 - 2 - 3 - 4
35. Dispersión	Dispersion	1 - 2 - 3 - 4
36. Distractores	Distracters	1 - 2 - 3 - 4
37. Distribución asimétrica	Biased distribution	1 - 2 - 3 - 4
38. Distribución bimodal	Bimodal distribution	1 - 2 - 3 - 4
39. Distribución de frecuencia	Frecuency distribution	1 - 2 - 3 - 4
40. Distribución normal	Normal distribution	1 - 2 - 3 - 4
41. Distribución de puntajes	Distribution scores	1 - 2 - 3 - 4
42. Distribución rectangular	Rectangular distribution	1 - 2 - 3 - 4
43. Edad mental	Mental age	1 - 2 - 3 - 4
44. Efecto de Halo	Halo effect	1 - 2 - 3 - 4
45. Efecto de práctica	Practice effect	1 - 2 - 3 - 4
46. Ejercicio de clasificación	Clasification exercise	1 - 2 - 3 - 4
47. Ejercicio de pareamiento	Matching excercise	1 - 2 - 3 - 4
48. Ejercicio de prueba	Test exercise	1 - 2 - 3 - 4

49. Enunciado del ítem	Item stem	1 - 2 - 3 - 4
50. Error de muestreo	Sampling error	1 - 2 - 3 - 4
51. Error estándar de medición	Standard error of measurement	1 - 2 - 3 - 4
52. Error de medición	Error of measurement	1 - 2 - 3 - 4
53. Error probable	Probable error	1 - 2 - 3 - 4
54. Error de varianza	Error variance	1 - 2 - 3 - 4
55. Escala	Scale	1 - 2 - 3 - 4
56. Escala de edad	Age scale	1 - 2 - 3 - 4
57. Escala de calidad	Quality scale	1 - 2 - 3 - 4
58. Escala gráfica de clasificación	Graphic rating scale	1 - 2 - 3 - 4
59. Estadístico	Statistic	1 - 2 - 3 - 4
60. Evaluación	Evaluation	1 - 2 - 3 - 4
61. Evaluación subjetiva	Subjective evaluation	1 - 2 - 3 - 4
62. Examen	Examination	1 - 2 - 3 - 4
63. Examen de ensayo	Essay examination	1 - 2 - 3 - 4
64. Examen externo	External examination	1 - 2 - 3 - 4
65. Escala de puntos	Point scale	1 - 2 - 3 - 4
66. Formas equivalentes	Equivalent forms	1 - 2 - 3 - 4
67. Fórmula de Kuder-Richardson	Kuder-Richardson formulas	1 - 2 - 3 - 4
68. Fórmula para calificar (cuando se restan % malas de buenas)	Scoring formula	1 - 2 - 3 - 4
69. Fórmula de Spearman-Brown	Spearman-Brown formula	1 - 2 - 3 - 4
70. Frecuencia acumulada	Cumulative frequency	1 - 2 - 3 - 4
71. Grupo de control	Control group	1 - 2 - 3 - 4

72. Heterogeneidad	Heterogeneity	1 - 2 - 3 - 4
73. Homogeneidad	Homogeneity	1 - 2 - 3 - 4
74. Incidente crítico	Critical incident	1 - 2 - 3 - 4
75. Índice de discriminación	Discrimination index	1 - 2 - 3 - 4
76. Inventario	Inventory	1 - 2 - 3 - 4
77. Índice de dificultad	Difficulty index	1 - 2 - 3 - 4
78. Ítem de falso-verdadero	True-false item	1 - 2 - 3 - 4
79. Ítem de mejor respuesta	Best answer item	1 - 2 - 3 - 4
80. Ítem de selección múltiple	Multiple choice item	1 - 2 - 3 - 4
81. Ítem de test	Test item	1 - 2 - 3 - 4
82. Máquina calificadora IBM	IBM scoring machine	1 - 2 - 3 - 4
83. Marca	Mark	1 - 2 - 3 - 4
84. Media	Mean	1 - 2 - 3 - 4
85. Mediana	Median	1 - 2 - 3 - 4
86. Medición	Measurement	1 - 2 - 3 - 4
87. Moda	Mode	1 - 2 - 3 - 4
88. Muestra aleatoria	Random sample	1 - 2 - 3 - 4
89. Muestra representativa	Representative sample	1 - 2 - 3 - 4
90. Norma	Norm	1 - 2 - 3 - 4
91. Norma de edad	Age norm	1 - 2 - 3 - 4
92. Norma de grado	Grade norm	1 - 2 - 3 - 4
93. Números aleatorios	Random numbers	1 - 2 - 3 - 4
94. Perfil	Perfile	1 - 2 - 3 - 4
95. Pertinencia	Relevance	1 - 2 - 3 - 4

96. Probabilidad	Probability	1 - 2 - 3 - 4
97. Promedio	Average	1 - 2 - 3 - 4
98. Prueba de rapidez	Speed test	1 - 2 - 3 - 4
99. Puntaje	Score	1 - 2 - 3 - 4
100. Puntaje burdo	Raw score	1 - 2 - 3 - 4
101. Puntaje compuesto	Composite score	1 - 2 - 3 - 4
102. Puntaje comparable	Comparable score	1 - 2 - 3 - 4
103. Puntaje crítico	Critical score	1 - 2 - 3 - 4
104. Puntaje derivado	Derived score	1 - 2 - 3 - 4
105. Puntaje estándar	Standard score	1 - 2 - 3 - 4
106. Puntaje estándar normalizado	Normalized standard score	1 - 2 - 3 - 4
107. Puntaje estaninos	Stanine scores	1 - 2 - 3 - 4
108. Puntaje de razón	Rate score	1 - 2 - 3 - 4
109. Puntaje T	T-score	1 - 2 - 3 - 4
110. Puntaje verdadero	True score	1 - 2 - 3 - 4
111. Puntajes z	z-scores	1 - 2 - 3 - 4
112. Pureza (de un test)	Purity (of a test)	1 - 2 - 3 - 4
113. Rango de puntajes	Range of scores	1 - 2 - 3 - 4
114. Rango percentil	Percentile rank	1 - 2 - 3 - 4
115. Rasgo, característica	Trait	1 - 2 - 3 - 4
116. Regresión	Regression	1 - 2 - 3 - 4
117. Sigma	Sigma	1 - 2 - 3 - 4
118. Sociometría	Sociometry	1 - 2 - 3 - 4
119. Sub-test	Sub-test	1 - 2 - 3 - 4
120. Sistema de honestidad	Honor system	1 - 2 - 3 - 4
121. Tabla de especificaciones	Table of specifications	1 - 2 - 3 - 4

122.	Tabla de expectación	Expectancy table	1 - 2 - 3 - 4
123.	Test	Test	1 - 2 - 3 - 4
124.	Test de analogías	Analogies test	1 - 2 - 3 - 4
125.	Test de aptitudes	Aptitudes test	1 - 2 - 3 - 4
126.	Test de calificación simultánea	Self-marking test	1 - 2 - 3 - 4
127.	Test de completación	Completion test	1 - 2 - 3 - 4
128.	Test de ejecución (de actuación)	Performance test	1 - 2 - 3 - 4
129.	Test de interpretación	Interpretive test	1 - 2 - 3 - 4
130.	Test de libro abierto	Open book test	1 - 2 - 3 - 4
131.	Test de maestría	Mastery test	1 - 2 - 3 - 4
132.	Test estandarizado	Standarized test	1 - 2 - 3 - 4
133.	Test de tiempo límite	Time limit test	1 - 2 - 3 - 4
134.	Test de diagnóstico	Diagnostic test	1 - 2 - 3 - 4
135.	Test no verbal	Non-verbal test	1 - 2 - 3 - 4
136.	Test libre de cultura	Culture free test	1 - 2 - 3 - 4
137.	Test de potencia	Power test	1 - 2 - 3 - 4
138.	Test proyectivo	Projective test	1 - 2 - 3 - 4
139.	Test objetivo	Objective test	1 - 2 - 3 - 4
140.	Test de rendimiento	Achievement test	1 - 2 - 3 - 4
141.	Test de Rorschach	Rorschach	1 - 2 - 3 - 4
142.	Técnica de selección forzada	Forced choice technique	1 - 2 - 3 - 4
143.	Test limitado por ejecución	Work limit test	1 - 2 - 3 - 4
144.	Validación cruzada	Cross validation	1 - 2 - 3 - 4
145.	Validez aparente	Face validity	1 - 2 - 3 - 4

146. Validez de contenido	Content validity	1 - 2 - 3 - 4
147. Validez estadística	Statistical validity	1 - 2 - 3 - 4
148. Validez	Validity	1 - 2 - 3 - 4
149. Varianza	Variance	1 - 2 - 3 - 4

ANEXO D

INVENTARIO DE MEDICIÓN DEL CONOCIMIENTO DE VOCABULARIO BÁSICO EN INVESTIGACIÓN EDUCATIVA

El siguiente inventario está constituido por 120 términos de uso en la investigación educativa y están dados en orden alfabético. Se da una escala de 4 puntos con el significado siguiente:

1. El término nunca lo ha encontrado en su experiencia en investigación educativa.
2. Ha escuchado del término en cuestión pero no sabe lo que significa.
3. Puede identificar el término en general pero no explicar su significado específico.
4. Puede explicar su significado específico

Tache su respuesta en la escala de 4 puntos

VOCABULARIO

1. Análisis de contenido	Content analysis	1 - 2 - 3 - 4
2. Bloque incompleto	Incomplete block	1 - 2 - 3 - 4
3. Comparación	Comparison	1 - 2 - 3 - 4
4. Comparación ortogonal	Orthogonal comparison	1 - 2 - 3 - 4
5. Confiabilidad	Reliability	1 - 2 - 3 - 4

6. Confiabilidad del observador	Observer reliability	1 - 2 - 3 - 4
7. Constructo hipotético	Hypothetical Construct	1 - 2 - 3 - 4
8. Correlación	Correlation	1 - 2 - 3 - 4
9. Cuadrados conjugados	Conjugate squares	1 - 2 - 3 - 4
10. Cuadrado estándar	Standard squares	1 - 2 - 3 - 4
11. Cuasiexperimento	Quasi-experiment	1 - 2 - 3 - 4
12. Cuestionario	Questionnaire	1 - 2 - 3 - 4
13. Datos básicos de referencia	Benchmark Data	1 - 2 - 3 - 4
14. Datos iniciales de referencia	Baseline Data	1 - 2 - 3 - 4
15. Definición operacional	Operational Definition	1 - 2 - 3 - 4
16. Desviación	Bias	1 - 2 - 3 - 4
17. Diseño de pretest solamente	Prettest-only design	1 - 2 - 3 - 4
18. Diseño de grupo paralelo	Parallel-group design	1 - 2 - 3 - 4
19. Diseño factorial	Factorial design	1 - 2 - 3 - 4
20. Diseño pretest-posttest	Prettest-posttest design	1 - 2 - 3 - 4
21. Distribución de muestras	Sampling distribution	1 - 2 - 3 - 4
22. Distribución normal multivariada	Multivariate normal distribution	1 - 2 - 3 - 4
23. Eficiencia relativa de un estadístico	Relative efficiency of a statistic	1 - 2 - 3 - 4
24. Encuesta	Survey	1 - 2 - 3 - 4
25. Entrevista	Interview	1 - 2 - 3 - 4
26. Entrevista estructurada	Structured interview	1 - 2 - 3 - 4

27. Entrevista no estructurada	Instructured interview	1 - 2 - 3 - 4
28. Error	Error	1 - 2 - 3 - 4
29. Error experimental	Experimental error	1 - 2 - 3 - 4
30. Error por azar	Chance error	1 - 2 - 3 - 4
31. Error sistemático	Systematic error	1 - 2 - 3 - 4
32. Error tipo I	Type I error	1 - 2 - 3 - 4
33. Error tipo II	Type II error	1 - 2 - 3 - 4
34. Escala graduada	Rating scale	1 - 2 - 3 - 4
35. Estadística	Statistics	1 - 2 - 3 - 4
36. Estimador	Estimator	1 - 2 - 3 - 4
37. Estimador consistente	Consistent estimator	1 - 2 - 3 - 4
38. Estimador eficiente	Efficient estimator	1 - 2 - 3 - 4
39. Estimador insesgado	Unbiased estimator	1 - 2 - 3 - 4
40. Evaluación por rangos	Rating	1 - 2 - 3 - 4
41. Fuente primaria	Primary source	1 - 2 - 3 - 4
42. Fuente secundaria	Secondary source	1 - 2 - 3 - 4
43. Generalización	Generalization	1 - 2 - 3 - 4
44. Grado de libertad	Degrees of freedom	1 - 2 - 3 - 4
45. Grupo de control	Control group	1 - 2 - 3 - 4
46. Hipótesis	Hypothesis	1 - 2 - 3 - 4
47. Hipótesis declarativa	Declarative hypothesis	1 - 2 - 3 - 4
48. Hipótesis de investigación	Research hypothesis	1 - 2 - 3 - 4
49. Hipótesis estadística	Statistical hypothesis	1 - 2 - 3 - 4
50. Hipótesis interrogativa	Question-form hypothesis	1 - 2 - 3 - 4

51. Hipótesis nula	Null hypothesis	1 - 2 - 3 - 4
52. Información relativa	Relative information	1 - 2 - 3 - 4
53. Interacción	Interaction	1 - 2 - 3 - 4
54. Intervalo de confianza	Confidence interval	1 - 2 - 3 - 4
55. Investigación aplicada	Applied research	1 - 2 - 3 - 4
56. Investigación básica	Basic research	1 - 2 - 3 - 4
57. Investigación de campo	Field research	1 - 2 - 3 - 4
58. Investigación de laboratorio	Laboratory research	1 - 2 - 3 - 4
59. Investigación descriptiva	Descriptive research	1 - 2 - 3 - 4
60. Investigación experimental	Experimental research	1 - 2 - 3 - 4
61. Investigación fundamental	Fundamental research	1 - 2 - 3 - 4
62. Investigación histórica	Historical research	1 - 2 - 3 - 4
63. Investigación pura	Pure research	1 - 2 - 3 - 4
64. Límites de confianza	Confidence limits	1 - 2 - 3 - 4
65. Matriz de varianza-covarianza	Variance-covariance matrix	1 - 2 - 3 - 4
66. Método de distribución libre	Distribution free method	1 - 2 - 3 - 4
67. Medición	Measurement	1 - 2 - 3 - 4
68. Método no paramétrico	Nonparametric method	1 - 2 - 3 - 4
69. Modelo de efectos aleatorios	Random-effects model	1 - 2 - 3 - 4
70. Modelo de efectos fijos	Fixed-effects model	1 - 2 - 3 - 4
71. Modelo estadístico	Statistical model	1 - 2 - 3 - 4
72. Modelo mixto	Mixed model	1 - 2 - 3 - 4
73. Muestra	Sample	1 - 2 - 3 - 4

74. Muestra aleatoria	Random sample	1 - 2 - 3 - 4
75. Nivel de significancia	Level of significance	1 - 2 - 3 - 4
76. Observación	Observation	1 - 2 - 3 - 4
77. Parámetro	Parameter	1 - 2 - 3 - 4
78. Pares combinados o apareamiento	Method pairs	1 - 2 - 3 - 4
79. Placebo	Placebo	1 - 2 - 3 - 4
80. Población	Population	1 - 2 - 3 - 4
81. Población hipotética	Hypotetical population	1 - 2 - 3 - 4
82. Postulado	Postulate	1 - 2 - 3 - 4
83. Pregunta abierta	Open question	1 - 2 - 3 - 4
84. Pregunta cerrada	Closed question	1 - 2 - 3 - 4
85. Procedimiento interfe- rente	Confounding	1 - 2 - 3 - 4
86. Probabilidad	Probability	1 - 2 - 3 - 4
87. Prueba estadística	Statistical test	1 - 2 - 3 - 4
88. Región crítica	Critical region	1 - 2 - 3 - 4
89. Región de aceptación	Acceptance region	1 - 2 - 3 - 4
90. Relación de causa-efecto	Cause-and-effect relation ship	1 - 2 - 3 - 4
91. Relación funcional	Functional relationship	1 - 2 - 3 - 4
92. Replicación	Replication	1 - 2 - 3 - 4
93. Selección al azar	Random assignment	1 - 2 - 3 - 4
94. Significación estadís- tica	Statistical significance	1 - 2 - 3 - 4
95. Significación práctica	Practical significance	1 - 2 - 3 - 4
96. Suma de cuadrados ajus- tados	Adjusted sum of squares	1 - 2 - 3 - 4

97. Teoría	Theory	1 - 2 - 3 - 4
98. Teoría de la decisión estadística	Statistical decision theory	1 - 2 - 3 - 4
99. Test de potencia	Test power	1 - 2 - 3 - 4
100. Test estadístico	Statistical test	1 - 2 - 3 - 4
101. Test de razón de verosimilitud	Likelihood ratio test	1 - 2 - 3 - 4
102. Test más poderoso	Most powerful test	1 - 2 - 3 - 4
103. Tratamiento	Treatment	1 - 2 - 3 - 4
104. Tratamientos cruzados	Crossed treatments	1 - 2 - 3 - 4
105. Tratamiento incluido	Nested treatment	1 - 2 - 3 - 4
106. Unidad de análisis	Unit of analysis	1 - 2 - 3 - 4
107. Unidad de muestreo	Sampling unit	1 - 2 - 3 - 4
108. Universo	Universe	1 - 2 - 3 - 4
109. Validez	Validity	1 - 2 - 3 - 4
110. Validez externa	External validity	1 - 2 - 3 - 4
111. Validez interna	Internal validity	1 - 2 - 3 - 4
112. Valor crítico	Critical value	1 - 2 - 3 - 4
113. Valor esperado	Expected value	1 - 2 - 3 - 4
114. Variable	Variable	1 - 2 - 3 - 4
115. Variable-estímulo	Stimulus variable	1 - 2 - 3 - 4
116. Variable independiente	Independent variable	1 - 2 - 3 - 4
117. Variable aleatoria	Random variable	1 - 2 - 3 - 4
118. Variable dependiente	Dependent variable	1 - 2 - 3 - 4
119. Variable interviniente	Intervening variable	1 - 2 - 3 - 4
120. Variable respuesta	Response variable	1 - 2 - 3 - 4

REFERENCIAS BIBLIOGRÁFICAS

- Ary, Donald. Introduction to Research in Education. New York: Holt, Rinehart and Winston, 1972.
- Best, John W. Como investigar en educación. Madrid: Morata, 1969.
- Cook, David R. A Guide to Educational Research. Boston: Allyn and Bacon, 1966.
- Fox, David J. The Research Process in Education. New York: Holt, Rinehart and Winston, 1969.
- Hayman, John L. Investigación y educación. Buenos Aires: Paidós, 1969.
- Strauss, Samuel. "Guidelines for Analysis of Research Reports". Journal of Educational Research, 1969, (63), págs. 165-169.
- Suydam, Marilyn N. "An Instrument for Evaluating Experimental Educational Research Reports". The Journal of Educational Research, 1968, (61), págs. 200-203.
- Zuñiga, Evenor, compilador. "La evaluación de las investigaciones en el ramo educativo". Investigaciones educativas en América. No. 2, 1967, (63), págs. 165-169.

